

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

MENU

SEARCH

INDEX

1/1



JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11)Publication number: 11027320

(43)Date of publication of application: 29.01.1999

(51)Int.Cl.

H04L 12/56
G06F 13/00
G06F 15/16
H04L 12/66

(21)Application number: 09178062

(71)Applicant:

FUJITSU LTD

(22)Date of filing: 03.07.1997

(72)Inventor:

IMAI YUJI

KISHIMOTO MITSUHIRO

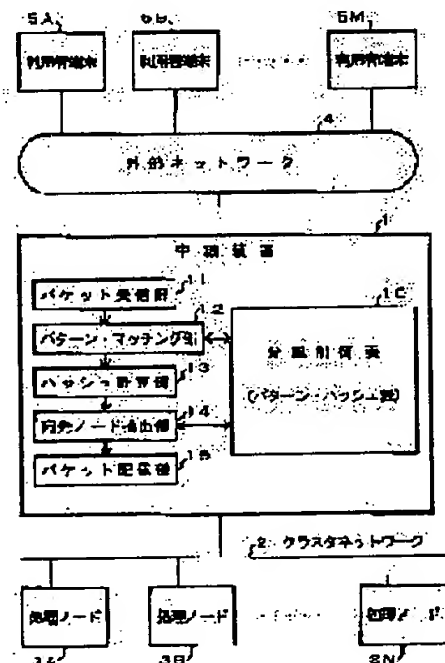
KATSUYAMA TSUNEO

(54) PACKET RELAY CONTROL METHOD, PACKET REPEATER AND PROGRAM STORAGE MEDIUM

(57)Abstract:

PROBLEM TO BE SOLVED: To simply and dynamically change packet distribution destinations and packet distribution conditions by realizing a single system image at a network address level in a control method that distributes packets to processing nodes in a device that connects an external network to a cluster network having plural processing nodes.

SOLUTION: Upon the receipt of a packet reaching an address representing a cluster from a packet reception section 11, a pattern matching section 12 of the repeater 1 uses a distribution control table 10 to conduct pattern matching based on a packet transmission source address, a source port and a destination port or the like. A hash calculation section 13 obtains an argument of a hash function based on the matching result to conduct hash calculation. A destination node extract section 14 decides a packet destination node based on the hash result by using the distribution control table 10.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-27320

(43) 公開日 平成11年(1999) 1月29日

(51) Int.Cl.⁶

識別記号

F I

H 0 4 L 12/56

H 0 4 L 11/20

1 0 2 D

G 0 6 F 13/00

3 5 5

G 0 6 F 13/00

3 5 5

15/16

3 7 0

15/16

3 7 0 N

H 0 4 L 12/66

H 0 4 L 11/20

B

審査請求 未請求 請求項の数18 O L (全 17 頁)

(21) 出願番号

特願平9-178062

(22) 出願日

平成9年(1997) 7月3日

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72) 発明者 今井 祐二

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(72) 発明者 岸本 光弘

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(72) 発明者 勝山 恒男

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(74) 代理人 弁理士 小笠原 吉義 (外2名)

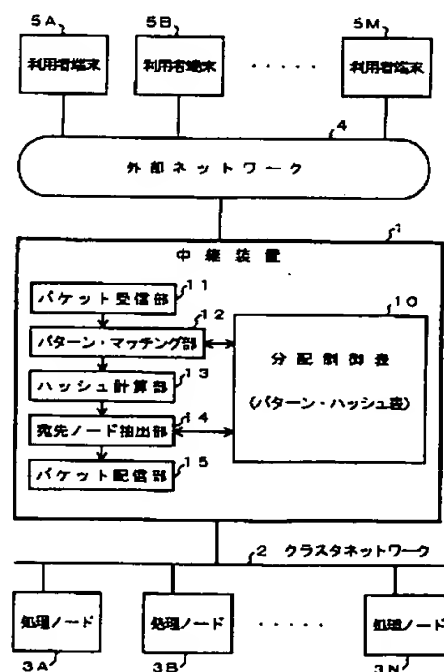
(54) 【発明の名称】 パケット中継制御方法、パケット中継装置およびプログラム記憶媒体

(57) 【要約】

【課題】外部ネットワークと複数の処理ノードを有するクラスタネットワークとを接続する装置においてパケットを処理ノードへ振り分ける制御方法に関し、ネットワークアドレスレベルでのシングルシステムイメージを実現し、かつパケットの振り分け先、振り分けの条件等を簡単かつ動的に変更できるようにすることを目的とする。

【解決手段】中継装置1のパターン・マッチング部12は、パケット受信部11からクラスタを代表するアドレスへ来るパケットを受け取ると、分配制御表10を用いて、そのパケットの発信元アドレス、発信元ポート、宛先ポート等をもとにパターン・マッチングを行う。ハッシュ計算部13は、マッチングの結果から、ハッシュ関数の引数を得てハッシュ計算を行う。宛先ノード抽出部14は、ハッシュ結果から分配制御表10を用いてパケットの宛先ノードを決定する。

本発明の構成例



【特許請求の範囲】

【請求項 1】 外部ネットワークと複数の処理ノードを有するクラスタネットワークとを接続する装置において、パケットを前記処理ノードへ振り分ける制御方法であって、パケットの発信元および宛先に関するいくつかのパターンに対してハッシュにより処理ノードを定める情報を記憶する分配制御表を持ち、前記外部ネットワークからクラスタを代表するアドレスへ来るパケットについて、前記分配制御表に定義されたパターンとのマッチングを行い、該当するパターンについて前記分配制御表により定まる処理ノードを選択し、選択した処理ノードへ前記パケットを配信することを特徴とするパケット中継制御方法。

【請求項 2】 請求項 1 記載のパケット中継制御方法において、前記パケットの発信元および宛先に関するパターンは、少なくとも発信元アドレス、発信元サービスエンドポイント指定子または宛先サービスエンドポイント指定子のいずれかを含む組み合わせのパターンであることを特徴とするパケット中継制御方法。

【請求項 3】 請求項 1 記載のパケット中継制御方法において、前記分配制御表は、パケットの発信元および宛先に関するパターンごとに、パターンの各々の構成要素をハッシュパラメータに使用するか使用しないかを指示する情報を持ち、ハッシュパラメータとして使用するパターン要素を抽出してハッシュを行うことにより前記分配制御表からパケットを配信する処理ノードを決定することを特徴とするパケット中継制御方法。

【請求項 4】 請求項 1 記載のパケット中継制御方法において、前記パターンごとまたはハッシュ項目ごとに、中継したパケットの数もしくは中継バイト数、またはその双方をカウントして記録し、記録したカウント情報を外部から参照可能としたことを特徴とするパケット中継制御方法。

【請求項 5】 請求項 1 記載のパケット中継制御方法において、前記分配制御表におけるパターンとハッシュ内容を、動作中に変更可能としたことを特徴とするパケット中継制御方法。

【請求項 6】 請求項 5 記載のパケット中継制御方法において、前記クラスタネットワークにおける処理ノードが故障したとき、故障した処理ノードへのハッシュ内容を他の故障していない処理ノードとなるように、前記分配制御表を動的に変更することを特徴とするパケット中継制御方法。

【請求項 7】 請求項 5 記載のパケット中継制御方法において、動作中に前記分配制御表を変更するときに、既存の VC を保存するための VC 保存パターンを設定可能としたことを特徴とするパケット中継制御方法。

【請求項 8】 請求項 7 記載のパケット中継制御方法において、コネクション切断確認パケットを監視し、コネクション切断確認パケットの検出により前記設定した V

C 保存パターンのエントリを自動消去することを特徴とするパケット中継制御方法。

【請求項 9】 請求項 7 記載のパケット中継制御方法において、コネクション切断確認パケットを監視しないで、前記処理ノードのサーバに現在接続中のコネクションを問い合わせ、現在接続中のコネクションがなくなっていることを確認して前記設定した VC 保存パターンのエントリを遅延消去することを特徴とするパケット中継制御方法。

10 【請求項 10】 請求項 1 記載のパケット中継制御方法において、前記外部ネットワークとクラスタネットワークとを接続する装置を、リクエストパケットを受信する装置と、返信用パケットを送信する装置とに分けたことを特徴とするパケット中継制御方法。

【請求項 11】 請求項 1 記載のパケット中継制御方法において、前記クラスタを代表するアドレスが複数あり、これらの代表アドレスへのパケットを一つ処理ノードに集約するように中継することを特徴とするパケット中継制御方法。

20 【請求項 12】 請求項 1 記載のパケット中継制御方法において、前記外部ネットワークとクラスタネットワークとを接続する装置が持つ前記分配制御表の状態を管理情報として、前記処理ノードの少なくとも一つに退避し、前記ネットワークを接続する装置の故障後のリスタート時、または故障時における他の装置への引き継ぎ時に、前記退避した管理情報により状態を回復することを特徴とするパケット中継制御方法。

30 【請求項 13】 外部ネットワークと複数の処理ノードを有するクラスタネットワークとを接続するシステムにおいてパケットを前記処理ノードへ振り分ける制御方法であって、前記クラスタネットワーク内の代表ノードは、パケットの発信元および宛先に関するいくつかのパターンに対してハッシュにより処理ノードを定める情報を記憶する分配制御表を持ち、前記外部ネットワークから前記代表ノードへ来るパケットについて、前記分配制御表に定義されたパターンとのマッチングを行い、該当するパターンについて前記分配制御表により定まる処理ノードを選択し、選択した処理ノードへ前記パケットを配信することを特徴とするパケット中継制御方法。

40 【請求項 14】 外部ネットワークと複数の処理ノードを有するクラスタネットワークとを接続する装置において、パケットの発信元および宛先に関するいくつかのパターンに対してハッシュにより処理ノードを定める分配制御表を記憶する手段と、前記外部ネットワークからクラスタを代表するアドレスへ来るパケットを受信する手段と、前記受信したパケットについて、前記分配制御表に定義されたパターンとのマッチングを行う手段と、該当するパターンにより定まるパラメータによってハッシュ計算を行う手段と、ハッシュ結果から前記分配制御表により定まる処理ノードを選択する手段と、選択した処

理ノードへ前記パケットを配信する手段とを備えたことを特徴とするパケット中継装置。

【請求項 15】 外部ネットワークと複数の処理ノードを有するクラスタネットワークとを接続する装置において実行するプログラムを記憶する媒体であって、前記外部ネットワークからクラスタを代表するアドレスへ来るパケットを受信する手段と、前記受信したパケットについて、あらかじめ記憶されたパケットの発信元および宛先に関するいくつかのパターンに対してハッシュにより処理ノードを定める分配制御表を用いて、この分配制御表に定義されたパターンとのマッチングを行う手段と、該当するパターンにより定まるパラメータによってハッシュ計算を行う手段と、ハッシュ結果から前記分配制御表により定まる処理ノードを選択する手段と、選択した処理ノードへ前記パケットを配信する手段とを実現するプログラムを格納したことを特徴とするプログラム記憶媒体。

【請求項 16】 外部ネットワークと複数の処理ノードを有するクラスタネットワークとを接続する装置におけるネットワークアドレスを変換する制御方法であって、パケットの発信元および宛先に関するいくつかのパターンに対して、ハッシュにより変換するネットワークアドレスを定めるパターン・ハッシュ表を持ち、前記処理ノードと前記外部ネットワークとの間で発着信するパケットについて、前記パターン・ハッシュ表に定義されたパターンとのマッチングを行い、パケットを発着信する処理ノードのネットワークアドレスを、該当するパターンによって前記パターン・ハッシュ表により定まるネットワークアドレスに変換することを特徴とするパケット中継制御方法。

【請求項 17】 外部ネットワークと複数の処理ノードを有するクラスタネットワークとを接続する装置において、パケットの発信元および宛先に関するいくつかのパターンに対してハッシュによりネットワークアドレスを定めるパターン・ハッシュ表を記憶する手段と、前記処理ノードと前記外部ネットワークとの間で中継するパケットを発着信する手段と、前記中継するパケットについて、前記パターン・ハッシュ表に定義されたパターンとのマッチングを行う手段と、該当するパターンにより定まるパラメータによってハッシュ計算を行う手段と、ハッシュ結果から前記パターン・ハッシュ表により定まる処理ノードのネットワークアドレスを抽出する手段と、パケットを発着信する処理ノードのネットワークアドレスを、該当するパターンによって前記パターン・ハッシュ表により定まるネットワークアドレスに変換する手段とを備えたことを特徴とするパケット中継装置。

【請求項 18】 外部ネットワークと複数の処理ノードを有するクラスタネットワークとを接続する装置において実行するプログラムを記憶する媒体であって、前記処理ノードと前記外部ネットワークとの間でパケットを中

継のために発着信する手段と、前記中継するパケットについて、あらかじめ記憶されたパケットの発信元および宛先に関するいくつかのパターンに対してハッシュによりネットワークアドレスを定めるパターン・ハッシュ表を用いて、このパターン・ハッシュ表に定義されたパターンとのマッチングを行う手段と、該当するパターンにより定まるパラメータによってハッシュ計算を行う手段と、ハッシュ結果から前記パターン・ハッシュ表により定まる処理ノードのネットワークアドレスを抽出する手段と、パケットを発着信する処理ノードのネットワークアドレスを、該当するパターンによって前記パターン・ハッシュ表により定まるネットワークアドレスに変換する手段とを実現するプログラムを格納したことを特徴とするプログラム記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は複数パケット交換方式ネットワークに接続される並列演算処理可能な計算機の通信機構に関し、特にアドレスレベルでのシングルシステムイメージを実現しつつ、クラスタ内ノードにネットワーク経由のリクエストパケットを分配し、負荷分散と高信頼化を実現可能としたパケット中継制御方法、パケット中継装置およびそれを実現するプログラムを格納したプログラム記憶媒体に関するものである。

【0002】 コンピュータネットワークのプロトコルとしては、ISO、CCITTによるOSIプロトコルスタック、TCP/IPなどが用いられる。これらのプロトコルスタックではトランスポート層のサービスとして、送信ホストの特定のプロセスから、受信ホストの特定プロセスへの通信を行う機構を定めている。

【0003】 トランスポートエンドポイントの指定は、ホスト指定とサービスエンドポイント指定子を組にして行われる。TCP/IPにおいては、ホスト指定がIPアドレス、サービスエンドポイント指定子がポート番号(port number)になる。受信計算機はこれを解析し、必要なプロセスに通信内容を伝送する処理を行う。

【0004】

【従来の技術】 大規模ネットワークサーバを構築するために、パーソナルコンピュータ(PC)やワークステーション(WS)を複数台接続するクラスタ構成を取ることがある。クラスタを構成する処理ノードは、ネットワークインタフェースを最低一つずつ持ち、それぞれにネットワークアドレスが割り当てられる。

【0005】 クライアントはクラスタにサービスを依頼する場合には、クラスタの中のどれか一つのノードのネットワークアドレスを獲得し、そこへリクエストパケットを送出する。通常、ネットワークアドレスの獲得は、ネームサーバへの問い合わせで行う。ネームサーバは、人間にとって意味のあるホスト名からネットワークアドレスへの変換データを持つデータベースを分散管理す

る。ある種のネームサーバが提供するネームサービスでは、クラスタが一つのシステムに見えるように、一つのホスト名に対して複数のネットワークアドレスを対応づけることができる。

【0006】例えば、インターネット標準のDNS (Domain Name System) は、これが可能である。DNSは、インターネットでドメイン名とIPアドレスの関係を管理する。クライアントは、複数のアドレスのうち一つを選択してリクエストする。通常、これがクライアント毎に異なるので、リクエストはクラスタ内のノードに分散する。

【0007】

【発明が解決しようとする課題】 サービスを受けるクライアントから見たときに、クラスタがあたかも高性能で故障しない1台のサーバであるかのように見えることが望ましい。このようにクラスタが一つのシステムとして見えるようにすることを、シングルシステムイメージ (Single System Image (SSI)) を持つと呼ぶ。SSIを持つことで、クライアントはサーバが単一ノードからできているときと同じようにクラスタにアクセスできる。

【0008】 複数のネットワークインタフェースを持つクラスタのSSIを実現する方法の一つに、ラウンド・ロビンDNS (round robin DNS) がある。図16は、本発明の課題を説明するための2台の処理ノードとIPアドレスを管理するノード(named) からなるhttpクラスタの例を示す図である。図中、4は外部ネットワーク、5A~5Mはサービスの提供を受ける利用者端末、9A~9Cは利用者端末5A~5Mに対してサービスを提供するノード、91はネットワーク間を接続するルータ(中継装置)、92はクラスタネットワークを表す。

【0009】 システムのネットワーク上での名前は、DNSを使用して登録・参照される。DNSでは、一つのドメイン名に複数のIPアドレスを登録することができる。利用者端末5A~5Mのクライアントに対して、ノード9A、9Bのhttpクラスタを一つの名前で見せるときには、このIPアドレスを管理するDNSの実現であるnamed ノード(ノード9C)に、ドメイン名www.domain.co.jpのIPアドレスレコード(Aレコード)として、httpdの動作するノード9A、9BのネットワークインタフェースのIPアドレスを、次の例のようにリストにして登録する。

【0010】 www.domain.co.jp IN A XXX.XXX.XXX.YYY ;: (Address of Node 9A)
www.domain.co.jp IN A XXX.XXX.XXX.ZZZ ;: (Address of Node 9B)

これを使用すると、クライアントから見てクラスタ全体を表す名前を代表としてアクセスすることができる。ここでは、ノード9Cは、クライアントからAレコードの

問い合わせがあるたびに、順番を並べ替えて答える。複数のクライアントがこのサーバにアクセスすると、それぞれ異なるノードのIPアドレスを検索結果として返信するので、結果的にリクエストは、各ノード9A、9Bのhttpdに確率的に分配される。

【0011】 通常、webクライアントを使用するユーザは、サーバをIPアドレスではなく、ドメイン名で認識するので、ドメイン名が一つのシステムイメージであれば、多くの場合十分である。この意味で、ラウンド・ロビンDNSは、ドメイン名レベルでSSIであるといえることができる。

【0012】 しかし、IPアドレスのレベルでは、依然としてノード9Aのアドレスと、ノード9Bのアドレスというように異なるアドレスが二つ存在し、一つのシステムのように見えてはいない。この点が問題を起こす場合がある。

【0013】 例えば、クラスタ内のノードの部分故障が発生したような場合である。ノード9Aとノード9Bが、同一のサービスを提供しているときに、ノード9Aが故障した場合を考える。DNSへの問い合わせで、たまたま故障したノード9AのIPアドレスを獲得してしまったクライアントは、ノード9Aからサービスを受けることができないので、httpdのサービスが停止しているように見える。連続して運転しているノード9Bがあっても、このクライアントがアドレスを獲得しなさい限りは、クライアントに対するサービスが回復しないことになる。

【0014】 本発明は上記問題点の解決を図り、アドレスレベルでのシングルシステムイメージ(SSI)を実現しつつ、クラスタ内ノードにネットワーク経由のリクエストパケットを分配し、高負荷リクエストに対応でき、かつ連続運転可能なネットワークサーバを実現することを目的とする。

【0015】

【課題を解決するための手段】 図1は、本発明の構成例を示す。1は中継装置、10は分配制御表(パターン・ハッシュ表ともいう)、11はパケット受信部、12はパターン・マッチング部、13はハッシュ計算部、14は宛先ノード抽出部、15はパケット配信部、2はクラスタネットワーク、3A~3Nは処理ノード、4は外部ネットワーク、5A~5Mは利用者端末を表す。

【0016】 中継装置1および処理ノード3A~3Nで、一つのクラスタを構成する。中継装置1は、分配制御表10、パケット受信部11、パターン・マッチング部12、ハッシュ計算部13、宛先ノード抽出部14、パケット配信部15を持ち、クラスタネットワーク2と外部ネットワーク4とを接続し、パケットを中継する。

【0017】 分配制御表(パターン・ハッシュ表)10は、パケットの発信元および宛先に関するいくつかのパターンに対してハッシュにより処理ノードを定めるため

の情報を記憶するものである。

【0018】パケット受信部11は、外部ネットワーク4からクラスタを代表するアドレスへ転送されてくるパケットを受信する手段である。パターン・マッチング部12は、受信したパケットについて、分配制御表10に定義されたパターンとのマッチングを行う手段である。

【0019】ハッシュ計算部13は、該当するパターンにより定まるパラメータによってハッシュ計算を行う手段である。宛先ノード抽出部14は、ハッシュ計算部13のハッシュ結果から分配制御表10により定まる処理ノードを選択する手段である。

【0020】パケット配信部15は、選択した処理ノード3A~3Nへ前記パケットを配信する手段である。以上の各処理手段を実現するプログラムは、計算機が読み取り可能な適当な記憶媒体に格納することができる。

【0021】本発明は、以下のように作用する。中継装置1のパケット受信部11は、外部ネットワーク4を介して、利用者端末5からのパケットを受け取ると、パターン・マッチング部12へ、そのパケットの例えばプロトコルタイプ、発信元アドレス、発信元サービスエンドポイント指定子、宛先サービスエンドポイント指定子を渡す。

【0022】パターン・マッチング部12は、分配制御表10を用いて、パターン・マッチングを行う。その結果を用いて、ハッシュ計算部13は、分配制御表10の構成に応じて定められた適当なハッシュ関数を用いてハッシュ計算を行い、宛先ノード抽出部14は、ハッシュ結果によって分配制御表10を参照し、パケットを配送の宛先ノードを決定する。

【0023】パケット配信部15は、宛先ノード抽出部14から通知されたノードへパケットを配送する。以上のように、パケットの特性に応じたパターンを用いたハッシュにより、分配制御表10によって、外部ネットワーク4からのリクエストを担当させる処理ノード3A~3Nの一つを決定する。したがって、ネットワークアドレスレベルでのシングルシステムイメージ（SSI）を実現することができる。特に、パターン・マッチングとハッシュとの組み合わせによって、負荷分散だけではなく、提供するサービス内容に応じた機能分散も同時に実現することが可能になる。

【0024】また、分配制御表10におけるパターンとハッシュの内容を動的に変更することにより、クラスタネットワーク2における処理ノード3A~3Nに部分故障が生じたような場合にも、シングルシステムイメージを継続したまま容易に対応することが可能になる。

【0025】

【発明の実施の形態】以下、本発明の実施の形態を説明する。ここでは、TCP/IPを対象にして説明するが、本発明はTCP/IPに限定されない。以降の説明では、ネットワークアドレス、サービスエンドポイント

などの用語の代わりに、具体例として、IPアドレス、ポートなどの一般的なインターネットの用語を用いる。

【0026】図1では、クラスタは中継装置1および処理ノード3A~3Nで構成されている。クラスタの処理ノード3A~3Nは、外部ネットワーク4に接続する利用者端末（クライアント）5A~5Mからは、仮想的なIPアドレス「X」を持つ1台のサーバとして認識される。処理ノード3A~3Nは、それぞれ独立にX宛のパケットを処理できるように、仮想IPアドレスとして「X」を持つ。

【0027】中継装置1を構成する計算機には、一般的なルーティングのプログラムを置き、外部ネットワーク4に対して仮想IPアドレスXの存在をアナウンスし、クラスタネットワーク2側に対して外部ネットワーク4のルート情報をアナウンスする。これにより、外部のネットワーク4のクライアント5A~5Mは、中継装置1を、サーバワークステーション（X）が接続されたネットワークへのルータであると認識することになる。

【0028】利用者端末5A~5Mから発信した仮想IPアドレスX宛のパケットは、ルーティング情報に従って中継装置1まで到達する。同時に、処理ノード3A~3Nから出た外部ネットワーク4向けパケットは、クラスタネットワーク2を通して中継装置1に向かって発信される。

【0029】このようなルーティングアナウンスのもとで、中継装置1は外部ネットワーク4から到着したパケットをプロトコルタイプprotocol type、発信元アドレスsrc address、発信元ポートsrc port、宛先ポートdest portを用いて選別し、処理ノード3A~3Nに振り分ける。選別は分配制御表（パターン・ハッシュ表）10を参照した簡単なパターン・マッチングとハッシュ計算により行う。

【0030】図2は、分配制御表の設定例を示している。分配制御表10は、図2に示すようにパターンテーブル10Pと、ノードテーブル10Nと、VC管理表10Vとからなる。パターンテーブル10Pは、パケットのパターンとそれらのパターンの各要素をハッシュパラメータに使用するかしないかを指示する情報、すなわちハッシュ関数に引数として渡すアドレス/ポートを指定する情報を持つテーブルである。ノードテーブル10Nは、ハッシュ結果をインデックスとする処理ノードのハッシュ表である。VC管理表10Vは、故障していた処理ノードをクラスタに復帰させるときにVCの状況を保存するためのテーブルである。

【0031】図2の設定例では、以下のような分配ポリシーを実現している。

1) 宛先ポート番号が80のhttp(80)へのリクエストは、発信元アドレスsrcaddrによりnode 0, node 1, node 2の各処理ノードへ分配して中継する。

【0032】2) 宛先ポート番号が8080のhttp(808

0)へのリクエストは、node 3の処理ノードへ中継する。
3)その他のリクエストは、発信元アドレスsrc addrをハッシュして、node 4, node 5, node 6の各処理ノードへ分配して中継する。

【0033】パターンテーブル10P中の「*」は、すべてのアドレス、ポートとマッチするパターンを表している。また、パターンテーブル10P中の「USE/NOT_USE」は、そのパターン要素をハッシュ関数のパラメータとして使用するかどうかを示している。

【0034】ハッシュ関数として、例えばハッシュパラメータの全パターンを1バイト単位で加算するものを用いれば、ハッシュ結果が8ビットであるので、ノードテーブル10Nのエントリは256個となる。4ビット単位で加算すれば、16個のエントリとなる。すべてが「NOT_USE」であれば、1個のエントリで処理ノードを決定することができるが、動的な変更を考慮するとエントリは多いほうが望ましい。ハッシュ関数をどのようなノードテーブル10Nに対しても汎用的に用いることができるようにするためには、前述したパターン要素の他にノードテーブル10Nのアドレスとサイズとを、引数で指定させるようにすればよい。

【0035】図示省略するが、中継装置1は、他に仮想IPアドレスリストを持つ。仮想IPアドレスリストは、中継装置1によるパケット分配を行う仮想IPアドレスのリストを保持する。

【0036】パターン・マッチング部12は、パケット受信部11(DLPIドライバ)から到着したパケットについて、まず、そのパケットの宛先IPアドレス(dest IP address)が仮想IPアドレスリストに載っているかどうかを検査する。リストに載っていればIPアドレスごとに指定されたパターンテーブル10Pの検査に進む。載っていない場合は、上位モジュールにそのままパケットを渡す。

【0037】仮想IPアドレスリストの検査で指定されたパターンテーブル10Pの先頭からマッチングを開始し、マッチした場合には指定されたパターン要素を用いたハッシュを行う。マッチしなかった場合には、順次、次のパターンテーブル10Pについてパターンの検査を行う。次のパターンテーブル10Pがなくなったのにマッチしない場合には、パケットは捨てられる。

【0038】TCPのパケットは、仮想回線(virtual circuit (VC))が同じならば、その宛先IPアドレス(dest IP address)、宛先ポート(dest port)、発信元ポート(src port)、発信元IPアドレス(src IP address)は同一である。パターン・マッチングは、宛先ポート(dest port)、発信元ポート(src port)、発信元IPアドレス(src IP address)を用いて行われるので、同じVCのパケットは、必ず同じパターンにマッチし、同じ処理ノードに配送される。

【0039】クラスタからクライアントに返送するパケ

ットには、IPヘッダの宛先アドレス(dest addr)にクライアントのIPアドレスが指定されており、ゲートする中継装置1のノードがアナウンスするルーティングによって、クラスタネットワーク2経由で中継装置1に向けて発信され、中継装置1により外部ネットワーク4へ流れていく。

【0040】振り分けポリシーは、前述したようにハッシュ関数と分配制御表10によって静的に実現する。TCP-VCセッションの開設・消滅とは無関係になるようにする。これは、以下の理由からである。

【0041】1)VCセッションの開設・消滅を中継装置1が意識する必要がない。もし中継装置1が、VCの開設・消滅を意識する必要があると、中継装置1はTCPの開設と消滅のリクエストパケットを監視しなければならない。これをクライアントとサーバの状態に不整合なく行うためには、論理が非常に複雑になり、かつ、オーバーヘッドも非常に大きくなる。

【0042】2)中継装置1自身の故障に対応するためには、故障直前の状態を保存し、故障を修復した後に状態を回復するか、もしくは処理を引き継ぐ別の中継装置に状態を回復する必要がある。計算に用いるハッシュ関数と分配制御表10が変化しないので、状態の変化も起こらず、状態の記録と保存、回復が容易になる。

【0043】パケットの振り分け処理について、さらに詳細に説明する。振り分け処理は、(1)パターン・マッチングとハッシュ法(Pattern match & Hash)、(2)VCプリザービング(VC preserving)の2種類のパターン・マッチングで行う。

【0044】最初に、本実施の形態で用いるパターン・マッチングとハッシュ法を説明する。この振り分け処理では、特定のパターンにマッチしたパケットを、指定された割合で各ノードにハッシュ関数を使用して分散する。

【0045】パターン・マッチングには、プロトコルタイプ(protocol type)、発信元アドレス(src addr)、発信元ポート(src port)、宛先ポート(dest port)を使用する。ハッシュ関数のパラメータは、発信元アドレス(src addr)、発信元ポート(src port)、宛先ポート(dest port)のうちから選択可能で、ハッシュ結果をインデックスとするノードの表を参照して、最終的にパケットを送付する処理ノードを決定する。

【0046】パターン・マッチングとハッシュ法を用いることで、サービスごとに処理ノードのグループを設定して負荷を分散させることができる。例えば、4ノード構成のクラスタ3A、3B、3C、3Dにおいて、4つの処理ノードを2:2に分割し、処理ノード3A、3Bを「http(80/TCP)サービス」専用に、処理ノード3C、3Dを「ftp(21/TCP)サービス」専用に割り当てるとする。

【0047】中継装置1が持つ分配制御表(パターン・

11

ハッシュ表) 10に、2つのパターン

(protocol type, src addr, src port, dest port) =
(TCP, *, *, 80)

(protocol type, src addr, src port, dest port) =
(TCP, *, *, 21)

と、パターンごとにハッシュ用の制御情報

(src addr, src port, dest port, table) = (USE, NOT-USE, NOT-USE, [1, 2])

(src addr, src port, dest port, table) = (USE, NOT-USE, NOT-USE, [3, 4])

とを設定する。

【0048】例えば、(TCP, *, *, 80)は、プロトコルタイプが「TCP」で、宛先ポートのポート番号が「80」のパケットのパターンとマッチする。発信元アドレスおよび発信元ポートは、パターンが「*」であるので、すべてのものとマッチする。一方、ハッシュ用の制御情報(USE, NOT-USE, NOT-USE, [1, 2])は、ハッシュパラメータとして発信元アドレスを用いたハッシュを行い、ハッシュ結果によって第1の処理ノード3Aまたは第2の処理ノード3Bのいずれかにパケットを分配することを示している。

【0049】すなわち、中継装置1に、例えばプロトコルタイプが「TCP」で、宛先ポートのポート番号が「80」であるhttp向けのパケットが到着すると、後述するVCプリザリングを行ったあと、上記パターンの検査を行う。この場合、1番目のパターンにマッチするので、上記ハッシュ用制御情報(ハッシュ表)の1番目を使いハッシュする。ハッシュ関数の引数とする発信元アドレス(src addr)はパケットのIPヘッダから値を取り出して指定し、発信元ポート(src port)、宛先ポート(dest port)には0を指定してハッシュ関数を実行することにより、処理ノード3A、3Bのどちらかを得る。したがって、http向けのパケットは発信元アドレス(src addr)によって処理ノード3A、3Bのどちらかのノードに配送され、処理されることになる。

【0050】代表的な振り分け方法の設定として以下のようなものがある。

1) 特定のsrc address からきたパケットを決まったノードに中継する。

pattern: (TCP, 111.222.33.44, *, *)

hash: (NOT-USE, NOT-USE, NOT-USE, [1])

2) 特定のdest port へのパケットを決まったノードへ中継する。

【0051】

pattern: (TCP, *, *, 80)

hash: (NOT-USE, NOT-USE, NOT-USE, [1])

3) src address をキーにしてハッシュを行い、n台のノードに分配して中継する。

【0052】

pattern: (TCP, *, *, *)

12

hash: (USE, NOT-USE, NOT-USE, [1, 2])

不特定多数のクライアントにサービスを提供する場合に、IPパケットのsrc address は多様になる。これをハッシュしてn台のノードの中から担当する処理ノードを決めることにより、リクエストを分配して処理することができるようになる。

【0053】4) src address とsrc portをキーにしてハッシュし、n台のノードに分配して中継する。

pattern: (TCP, *, *, *)

10 hash: (USE, USE, NOT-USE, [1, 2])

上記3)で十分に負荷分散ができないほどクライアントが少ない場合に、ハッシュのキーとしてsrc portを含める。TCPではVCの一意性を確保するためにVCごとにsrc portを変えており、これにより、同じクライアントからの別のVCを異なるノードでサービスすることができる。

【0054】5) 特定のdest port へのパケットを、src address のハッシュを用いてn台のノードに分配して中継する。

20 pattern: (TCP, *, *, 80)

hash: (USE, NOT-USE, NOT-USE, [1, 2])

プロトコル毎に専用サーバをn台用意し、リクエストをn台に分散する。

【0055】これ以外にも様々に組み合わせて使用することができる。このように、分配制御表10の設定を任意に変更し操作することにより、振り分け先の処理ノード、振り分けの割合、振り分けるパケットの種類を容易に、また、必要に応じて動的に制御することが可能になる。

30 【0056】図3～図6を用いて、中継装置の基本動作を具体例に従ってさらに詳しく説明する。図3に示すような、クライアントおよびクラスタの構成において、クライアント51～53のIPアドレスをそれぞれ「A、B、C」、サーバ側のノード31～33のIPアドレスをそれぞれ「I、J、K」、クラスタの代表IPアドレスを「X」とする。ノード31～33は、「httpd」、

「nntpd」により、仮想IPアドレスXでハイパーテキストを使ったインターネットの情報サービスhttp(80)と、ネットニュースのメッセージ転送サービスnntp(119)を提供している。ノード31～33は、リクエストの処理が1:1:2になるように調整するものとする。

【0057】この場合、中継装置1の持つ分配制御表(パターン・ハッシュ表)10は、図4に示すように設定され、ノード31～33のMACアドレス表は、図5に示すように設定される。

【0058】図6は、パケット受信時の中継装置1の処理のフローチャートである。図6において、ステップS1では、パケットの宛先が代表IPアドレスかどうかを判断する。宛先が代表IPアドレスであればステップS2へ進み、そうでなければステップS9へ進む。

【0059】ステップS2では、代表IPアドレスに対して設定された図4に示すパターン・ハッシュ表からマッチするパターンを探す。ステップS3の判定により、マッチするパターンがある場合にはステップS4へ進み、マッチするパターンがない場合にはステップS9の処理を行う。

【0060】ステップS4では、必要なハッシュパラメータを用いてハッシングする。すなわち、ヘッダから発信元アドレス、発信元ポート、宛先ポートを取り出して、このうちパターン・ハッシュ表でUSEになっているものを用いてハッシュ計算を行う。図4のパターン・ハッシュ表の例では、発信元IPアドレス (srcIP) を用いてハッシュ計算を行うことになる。

【0061】ステップS5では、ハッシュ結果から宛先ノード番号を得る。図4に示すパターン・ハッシュ表のresult番目から宛先ノード番号を得る。例えば、ハッシュ結果が65であれば、宛先ノード番号は2となる。

【0062】ステップS6では、宛先ノード番号からMACアドレスを得る。図5のMACアドレス表を用いた場合、宛先ノード（ノード番号=2）からMACアドレス「j」を得る。

【0063】ステップS7では、IPパケットのMACアドレスを付け替える。ここでは、IPアドレスのXをjに書き換える。ステップS8では、クラスタネットワーク2へパケットを送信し、中継処理を終了する。

【0064】ステップS9では、分配先が定義されていないので、そのパケットを廃棄する。次に、故障発生時、故障回復時の動作について説明する。

【0065】ノードの部分故障は、上位の管理マネージャにより検出される。上位の管理マネージャは、故障前の分配の設定から、分配制御表10に故障したノードテーブルが載っていれば、それを代替ノードに置き替えて、ノードテーブルを再登録する。

【0066】引き継ぎ後も、負荷をできるだけ均等にするためには、あらかじめノード数よりノードテーブルの大きさを多めに作っておき、同じノードを複数のハッシュエントリに対して登録しておくようにする。故障ノード以外は、経路が変わらないので、故障ノード以外で処理中のVCは継続される。

【0067】クラスタ内のノードのうち、例えばノード32が故障した場合には、中継装置1の内部の分配制御表を、図7に示すように変更し、ノード番号2に割り当てられていた範囲をノード番号1または3に書き換えることにより、故障したノード32に割り当てていた範囲を故障していない他のノード31、33に割り当て直す。その際、非故障ノードの割合を崩さないように配分を行う。

【0068】次に、故障して論理的に切り離していたノードを、修復後にクラスタに復帰させる場合について説明する。例えば、保守のために切り離していたノードを

復帰させる場合も同様である。

【0069】例えば図8に示すようなクラスタ構成において、ノード31が故障し、その後、修復されたとする。ノード31が故障から回復するときに、ノード32、ノード33には、図8に示すようなVCが開設されているとする。

【0070】クライアント52からノード32へのVCは、(srcIP, srcPort, dstPort, destIP) = (B, 2468, 119, X)、クライアント53からノード33へのVCは (srcIP, srcPort, dstPort, destIP) = (C, 1234, 80, X) である。

【0071】故障中だったノードが回復し、クラスタに復帰する場合には、変更したノードテーブルの該当する項目を単純に書き直して再設定すると、そのノードで処理中のVCが切断されてしまう。これを回避するために、再設定直前のVCの状態を保存する。すなわち、VCプリザービング (VC preserving) を行う。ここで、特定のVCを特定のノードに送付すると、VCは、発信元アドレス、発信元ポート、宛先ポート、宛先アドレスの4つの項目の組で特定される。中継装置1は、このVCをこれらの4つの項目の組で設定して記憶し、これと対応させて宛先ノードを設定しておく。パケットが到着するとパターン・マッチングを行い、合致したパケットを指定されたノードに配送する。

【0072】VCプリザービングは、ホットスイッチ (hot switch) 時のTCP__VCを保存するために使用される。ホットスイッチとは、故障していたノードをクラスタに復帰させることである。このときに、新規のVCを回復ノードに収容するために、分配制御表10におけるノードテーブルの設定を変更するが、この際、引き継いでいたノードにおいてサービス中であったVCは、切断されてしまわないように保存する必要がある。

【0073】図8に示すような (srcIP, srcPort, dstPort, destIP) = (B, 2468, 119, X) のVCが、新たなパターン・ハッシュ表でノード31に割り当てられる場合には、図9に示すようなエントリをVC管理表に追加する。

【0074】図10は、故障したノードをクラスタに復帰させるときのVC管理表作成処理のフローチャートである。ステップS11では、すべての処理ノードについて、以下のステップS12～S13の処理を繰り返す。

【0075】ステップS12では、現在開設中のVCリストを獲得する。ステップS13では、新旧パターン・ハッシュ表を検索し、ノード宛先をそれぞれ求める。

【0076】ステップS14では、新旧パターン・ハッシュ表でノードの変更があるかどうかを判断する。ノードの変更がある場合にはステップS15の処理を行い、ノードの変更がない場合にはステップS16へ進む。

【0077】ステップS15では、VCを旧パターン・ハッシュ表でのVC宛先とともにVC管理表に登録する。ステップS16では、すべての処理ノードについて

処理が終了したかどうかを判定し、終了するまでステップS 1.1へ戻って処理を繰り返す。すべての処理ノードに対する処理が終了した場合には処理を終了する。

【0078】管理マネージャは、引き継ぎ完了後、定期的に引き継いだノードに対して、VCが終了したかどうかを問い合わせて検査する。終了しているVCがあれば、VC管理表の該当するVCブリザーブエントリを削除する。なお、コネクション切断確認パケット(FINパケット)を監視して、コネクション切断確認パケットの検出により、VC管理表の該当するVCブリザーブエントリを自動消去するようにしてもよい。

【0079】図11は、VC管理表を用いる場合のパケット受信時の処理のフローチャートである。ステップS 21では、パケットの宛先は代表IPアドレスかどうかを判断する。宛先が代表IPアドレスの場合にはステップS 22へ進み、そうでない場合にはステップS 32へ進む。

【0080】ステップS 22では、VC管理表の各エントリに着目し、VCがマッチするかどうかを判断する。VCがマッチする場合にはステップS 23へ進み、マッチしない場合にはステップS 25へ進む。

【0081】ステップS 23では、VC管理表のVCにマッチした場合、VC管理表に従って宛先MACアドレスを書き換える。ステップS 24では、クラスタネットワーク2へパケットを送信して、中継処理を終了する。

【0082】VC管理表のVCにマッチしない場合、ステップS 25では、処理ノードへのパケットの振り分けのため、パターン・ハッシュ表からマッチするパターンを探す。

【0083】ステップS 26の判定により、マッチするパターンがあればステップS 27へ進み、マッチしない場合にはステップS 32へ進む。ステップS 27では、必要なハッシュパラメータを用いてハッシングする。

【0084】ステップS 28では、ハッシュ結果からノードテーブルを参照し宛先ノード番号を得る。ステップS 29では、宛先ノード番号からMACアドレスを得る。

【0085】ステップS 30では、IPパケットのMACアドレスを付け替える。ステップS 31では、クラスタネットワークへパケットを送信し、中継処理を終了する。

【0086】ステップS 32では、分配先が定義されていないので、そのパケットを廃棄する。次に、中継装置自身の故障等に対処するため、中継装置を二重化した場合の例を説明する。この場合には、管理情報引き継ぎのためのパターン・ハッシュ表の変更時の処理および再立ち上げ時の処理を図12に示すようにして行う。図12(A)は、パターン・ハッシュ表の変更時の処理のフローチャート、図12(B)は、再立ち上げ時の処理のフローチャートである。

【0087】図12(A)において、ステップS 41では、パターン・ハッシュ表の変更があったかどうかを判断する。パターン・ハッシュ表の変更があった場合にはステップS 42へ進み、変更がなかった場合には処理を終了する。

【0088】ステップS 42では、パターン・ハッシュ表の情報をローカル・ディスクに格納する。ステップS 43では、中継装置が二重化構成であるかどうかを判断する。中継装置が二重化構成である場合には、ステップS 44へ進み、二重化構成でない場合には処理を終了する。

【0089】ステップS 44では、引き継ぎ装置のファイルにパターン・ハッシュ表の情報を格納する。図12(B)の再立ち上げ時の処理において、ステップS 51では、パターン・ハッシュ表の情報をローカル・ディスクから読み込む。

【0090】ステップS 52では、処理ノードからVCの状態を獲得する。ステップS 53では、すべてのVC状態について、以下のステップの処理を行う。

【0091】ステップS 54では、各VCの状態についてパターン・ハッシュ表に適用した時の宛先ノードを計算する。ステップS 55では、パターン・ハッシュ表に適用した時に、実際の宛先ノードと異なるかどうかを判断する。実際の宛先ノードと異なる場合には、ステップS 56へ進み、実際の宛先ノードと異なる場合にはステップS 57へ進む。

【0092】ステップS 56では、VCの状態をVC管理表に登録する。ステップS 57では、すべてのVCについての処理が終了したかどうかを判断し、終了している場合にはステップS 58へ進み、終了していない場合にはステップS 53以降を同様に繰り返す。

【0093】ステップS 58では、実際に用いるパターン・ハッシュ表を変更する。中継装置1を二重化する場合に、負荷分散および機能分散を考慮して、リクエストパケットを受信する装置と、返信用パケットを送信する装置とに分けて構成しておき、一方が故障したときに、その機能を他方の装置が引き継ぐように構成するようにしてもよい。

【0094】さらに、本実施の形態では、複数アドレスを1ノードに集約してパケットを配信することも可能である。図13は、複数アドレスを1ノードへ集約する場合の例を示す。

【0095】例えば、図13(A)に示すように、一つのノード31で、ハイパーテキストを使ったインターネットの情報サービスへのリクエストhttp(80)と、ネットニュースのメッセージ転送サービスへのリクエストnnlp(119)を受け取るが、それぞれXとYの複数のIPアドレスで受け取りたいとする。この場合、パターン・ハッシュ表を図13(B)に示すように設定する。このように、異なる複数の代表IPアドレスのパターン・ハッシュ

ュ表を同様の内容で設定することにより、一つのノードへ複数アドレスを集約することが可能となる。

【0096】以上の実施の形態において、図2に示すパターンテーブル10Pのパターンごと、またはノードテーブル10Nのノードごとに、中継したパケットの数および中継バイト数をカウントして記録しておき、この記録したカウント情報を処理ノードの一つである管理ノードから参照可能とすれば、このカウント情報を利用して、処理ノードの負荷の均一化を図るための分配制御表10の再設定を容易に行うことが可能になる。

【0097】図1.4は、本発明の他の構成例を示す。図1に示す例では、中継装置1が分配制御表10を持ち、パターン・マッチングとハッシュにより、クラスタの代表アドレスへ来たパケットを各処理ノード3A~3Nに振り分けるようにしていた。図1.4の例では、外部ネットワーク4とクラスタネットワーク2とを接続する中継装置100には、通常のルータの機能だけを持たせ、処理ノード群の中の一つを代表の処理ノード30として、これにパケットの分配機能を持たせる。

【0098】処理ノード30は、前述した中継装置1が持つ分配制御表10、パケット受信部11、パターン・マッチング部12、ハッシュ計算部13、宛先ノード抽出部14、パケット配信部15に相当する手段を持ち、分配制御表10を用いたパターン・マッチングを行った後、分配制御表10におけるノードテーブル（ハッシュテーブル）に従って、各処理ノード31~33にリクエストパケットを振り分ける。これによっても同様の機能を実現することができる。

【0099】以上の実施の形態では、外部ネットワーク4からきたパケットを各処理ノードへ振り分ける例を説明したが、クラスタ側から外部へパケットを発信する場合に、同様な原理によってネットワークアドレスを変換して、クラスタ内のネットワークアドレスの隠蔽を図ることも可能である。

【0100】図1.5は、クラスタ内のネットワークアドレスを変換して外部へパケットを発信するシステムの構成例を示す。中継装置1は、図1.5(A)に示すようにパターン・ハッシュ表10'とネットワークアドレス変換部16を持つ。処理ノード3A~3Nの一つが外部ネットワーク4に対してパケットを発信するとき、ネットワークアドレス変換部16は、パケット中のIPヘッダの発信元アドレスを、パターン・ハッシュ表10'を用いて他の発信元アドレスに書き換える。例えば、処理ノード3Aと処理ノード3Bのアドレスを共通の発信元アドレスに書き換えることにより、外部からは処理ノード3Aと処理ノード3Bとを一つのノードに見せることが可能になる。

【0101】パターン・ハッシュ表10'は、図1.5(B)に示すようにパターンテーブル10Pとハッシュテーブル10Hとからなる。パターンテーブル10Pの

構成は、図2に示したものと同様である。ハッシュテーブル10Hの構成は、ノードテーブル10Nと同様であるが、各ハッシュエントリに宛先ノードの情報ではなく、変換する新しい発信元アドレスを設定する。ネットワークアドレス変換部16におけるパターン・マッチングとハッシュの処理は、前述したリクエストパケットに対するときの処理とまったく同様である。

【0102】

【発明の効果】以上説明したように、本発明によれば、ネットワークアドレスレベルでのシングルシステムイメージを実現し、パターン・マッチングとハッシュによりクラスタ内の宛先ノードを決めてパケットを配信することができる。特に、パターン・ハッシュ表に設定する情報によって、パケットの振り分け条件を任意に設定することが可能である。さらに、パターン・ハッシュ表を動的に変更することにより、パケットの種類やノードの処理負荷分散の状況をみて、パケットの振り分け先を容易に変更することも可能である。

【0103】また、VC管理表により、VCの状態を保存することができ、故障等の場合にも、二重化構成による引き継ぎや復旧を問題なく簡単に行うことができる。このように、クラスタネットワークリクエストのアドレスパターンによる高速負荷分散、およびノード故障時、保守時などのノードの部分停止の場合の連続運転が可能になる。

【図面の簡単な説明】

【図1】本発明の構成例を示す図である。

【図2】分配制御表（パターン・ハッシュ表）の設定例を示す図である。

【図3】実施の形態を説明するためのクラスタの構成例を示す図である。

【図4】分配制御表の構成例を示す図である。

【図5】MACアドレス表の例を示す図である。

【図6】パケット受信時の処理のフローチャートである。

【図7】分配制御表の構成例を示す図である。

【図8】VC管理表の構成例を示す図である。

【図9】分配制御表の構成例を示す図である。

【図10】VC管理表作成処理のフローチャートである。

【図11】VC管理表を用いる場合のパケット受信時の処理のフローチャートである。

【図12】管理情報引き継ぎ処理のフローチャートである。

【図13】複数アドレスを1ノードへ集約する場合の例を示す図である。

【図14】本発明の他の構成例を示す図である。

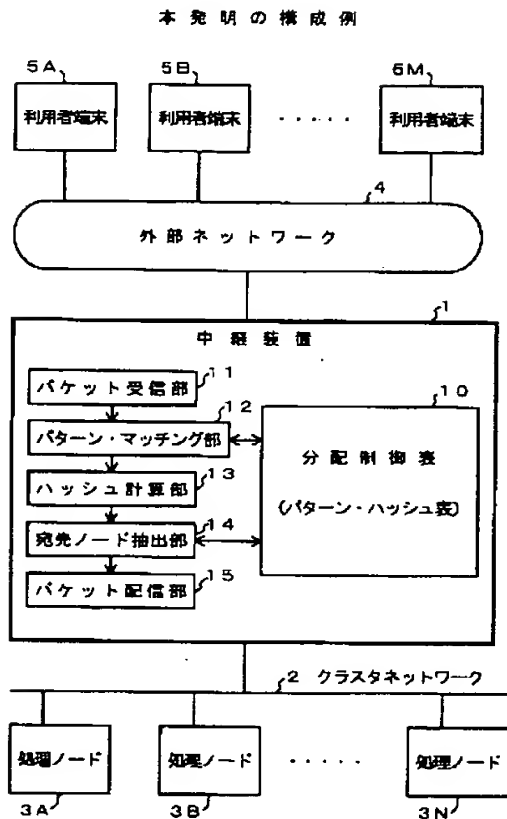
【図15】ネットワークアドレスを変換して外部へパケットを発信するシステムの構成例を示す図である。

【図16】本発明の課題を説明するための図である。

【符号の説明】

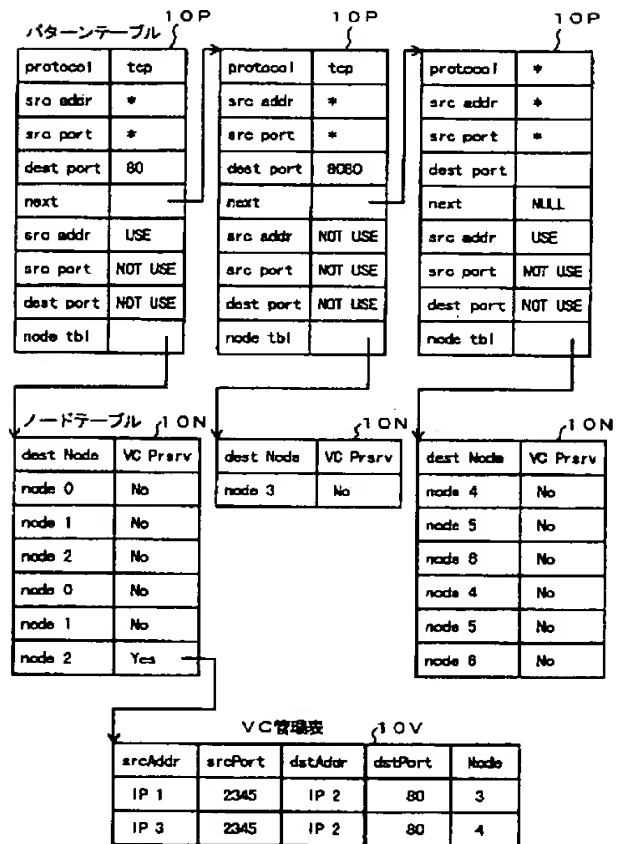
- 1 中継装置
2 クラスタネットワーク
3 A ~ 3 N 処理ノード
4 外部ネットワーク
5 A ~ 5 M 利用者端末

【図 1】



- 1 0 分配制御表 (パターン・ハッシュ表)
1 1 パケット受信部
1 2 パターン・マッチング部
1 3 ハッシュ計算部
1 4 宛先ノード抽出部
1 5 パケット配信部

【図 2】



【図 5】

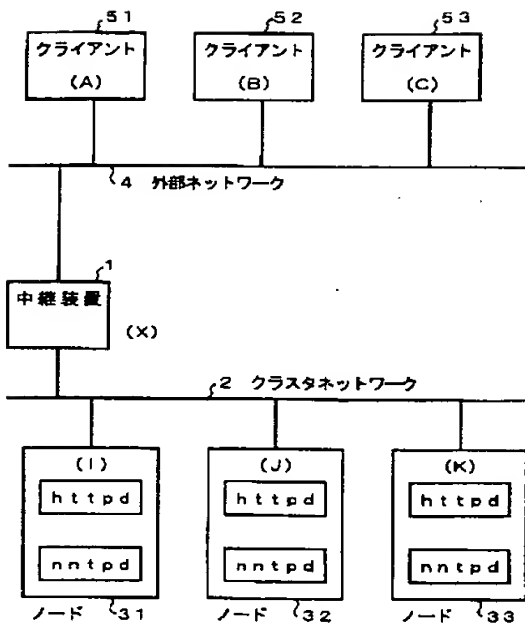
MACアドレス表の例

| ノード | MAC addr |
|-----|----------|
| 1 | i |
| 2 | j |
| 3 | k |

【図 9】

| VC | | | | 宛先ノード |
|-------|---------|--------|----------|-------|
| srcIP | srcPort | destIP | destPort | |
| B | 2468 | X | 119 | ノード32 |

【図 3】



【図 4】

分配制御表 (パターン・ハッシュ表) の構成例

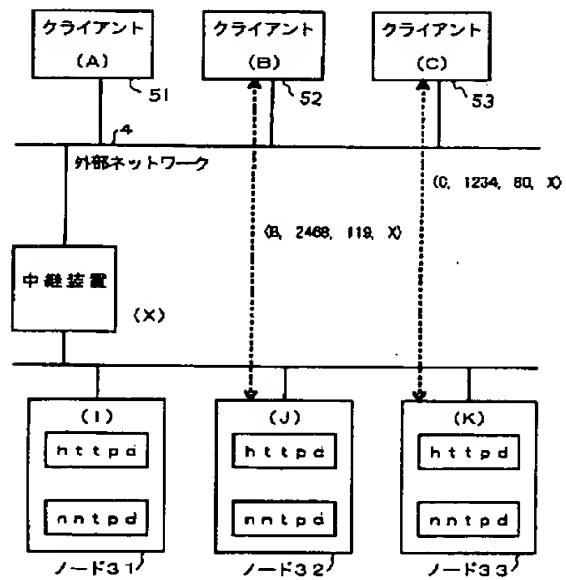
| 代表 IP | パターン | | | | ノード | |
|-------|-----------|-------|----------|-------------|------------|----|
| | protoType | srcIP | srcPort | destPort | result | 宛先 |
| X | TOP | *USE | *NOT-USE | 80/NOT-USE | size = 256 | |
| | | | | | 0 ↓ 1 | 1 |
| | | | | | 63 ↓ 1 | 1 |
| | | | | | 64 ↓ 2 | 2 |
| | | | | | 127 ↓ 2 | 2 |
| | | | | | 128 ↓ 3 | 3 |
| | | | | | 255 ↓ 3 | 3 |
| | TOP | *USE | *NOT-USE | 119/NOT-USE | size = 256 | |
| | | | | | 0 ↓ 1 | 1 |
| | | | | | 63 ↓ 1 | 1 |
| | | | | | 64 ↓ 2 | 2 |
| | | | | | 127 ↓ 2 | 2 |
| | | | | | 128 ↓ 3 | 3 |
| | | | | | 255 ↓ 3 | 3 |

【図 7】

分配制御表 (パターン・ハッシュ表) の構成例

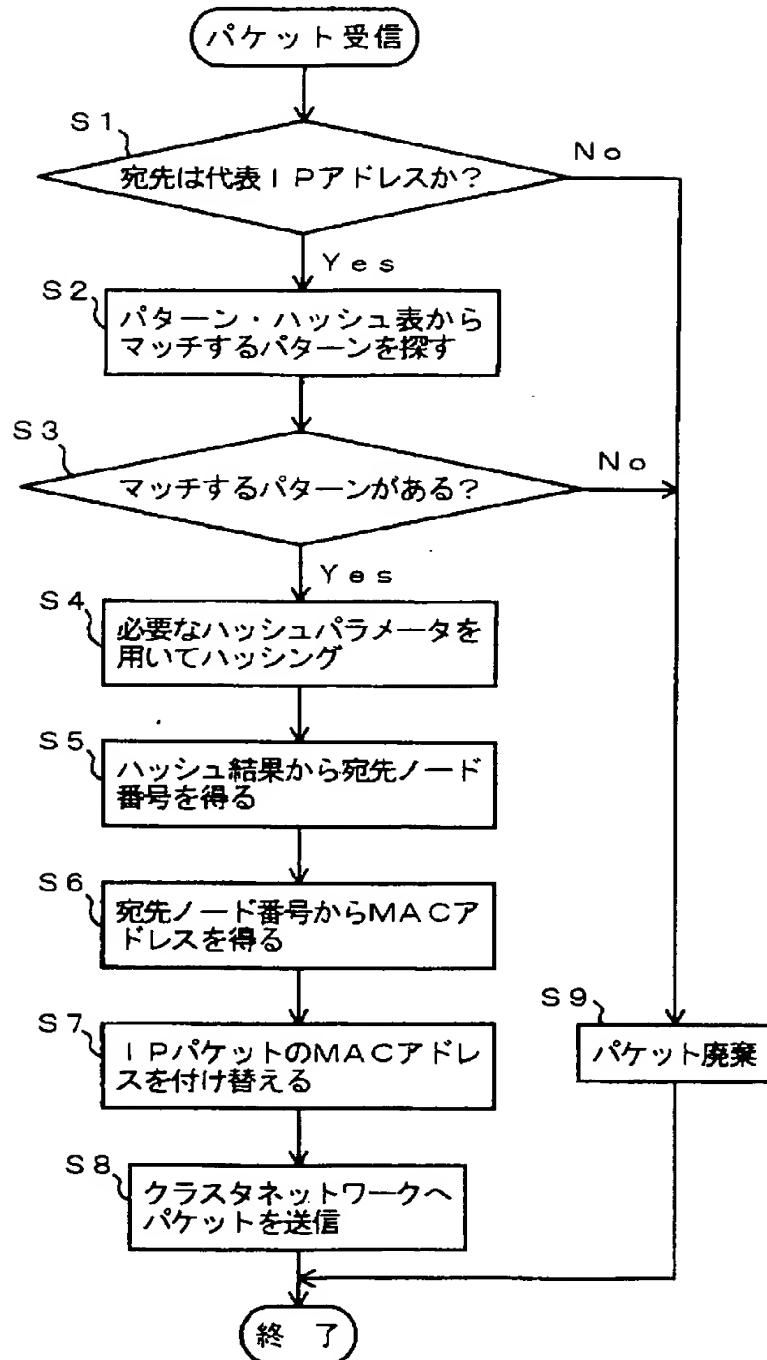
| 代表 IP | パターン | | | | ノード | |
|-------|-----------|-------|----------|-------------|------------|----|
| | protoType | srcIP | srcPort | destPort | result | 宛先 |
| X | TOP | *USE | *NOT-USE | 80/NOT-USE | size = 256 | |
| | | | | | 0 ↓ 1 | 1 |
| | | | | | 63 ↓ 1 | 1 |
| | | | | | 64 ↓ 2 | 2 |
| | | | | | 127 ↓ 2 | 2 |
| | | | | | 128 ↓ 3 | 3 |
| | | | | | 255 ↓ 3 | 3 |
| | TOP | *USE | *NOT-USE | 119/NOT-USE | size = 256 | |
| | | | | | 0 ↓ 1 | 1 |
| | | | | | 63 ↓ 1 | 1 |
| | | | | | 64 ↓ 2 | 2 |
| | | | | | 127 ↓ 2 | 2 |
| | | | | | 128 ↓ 3 | 3 |
| | | | | | 255 ↓ 3 | 3 |

【図 8】

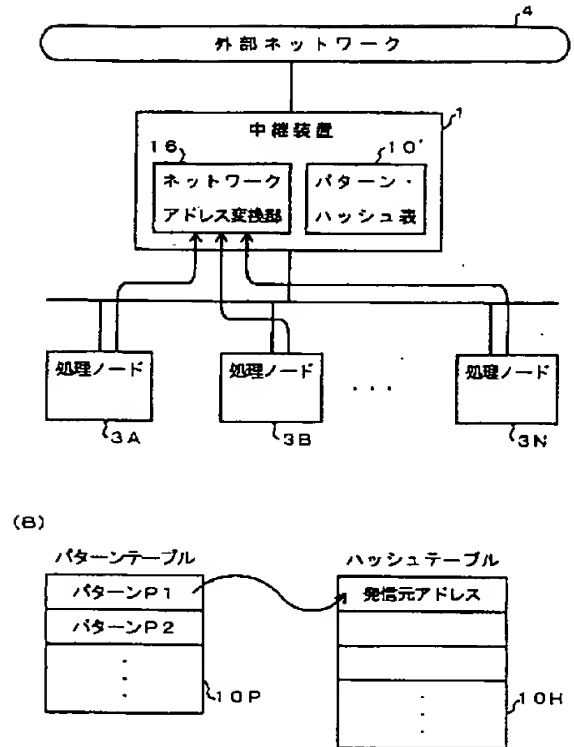


【図6】

パケット受信時の処理のフローチャート

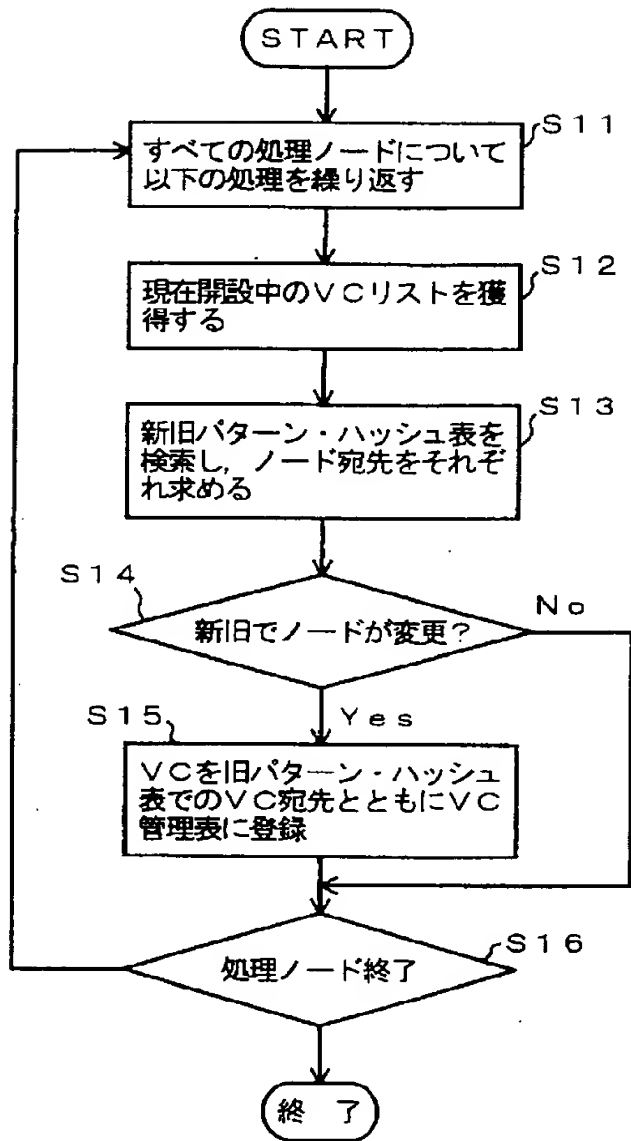


【図15】



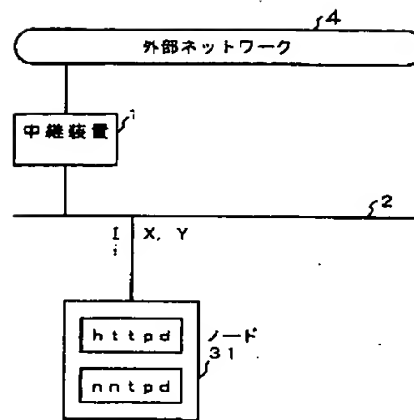
【図10】

VC管理表作成処理のフローチャート



【図13】

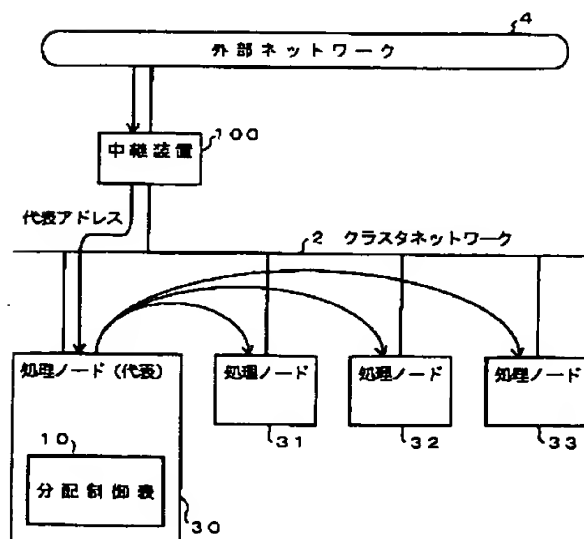
(A)



(B)

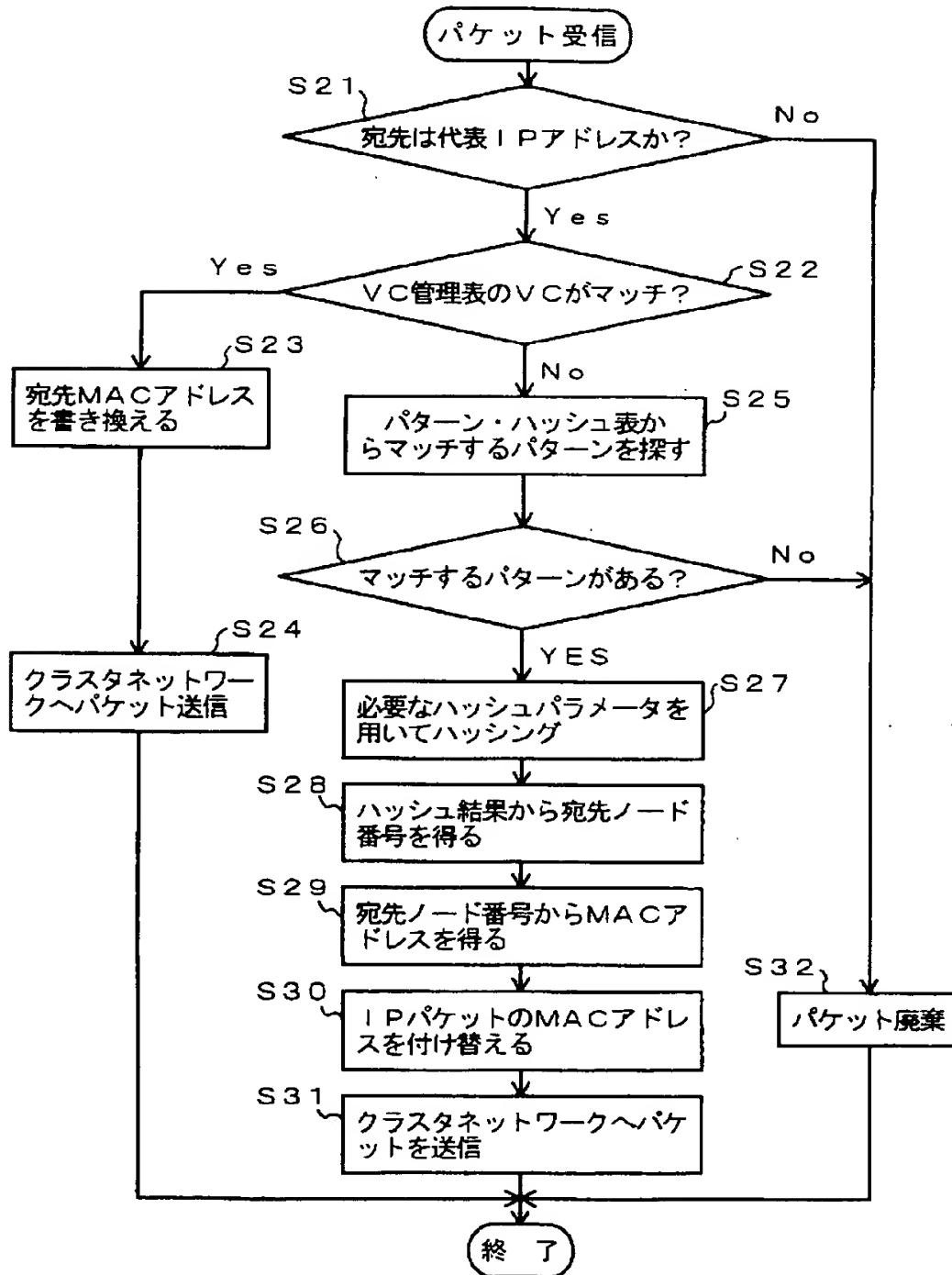
| 代表IP | パターン | | | | ノード |
|------|-----------|-------|----------|-------------|-----------------|
| | protoType | srcIP | srcPort | destPort | |
| X | TCP | *USE | *NOT-USE | 80/NOT-USE | size = 1 0 1 |
| Y | TCP | *USE | *NOT-USE | 119/NOT-USE | size = 1 0 1 |

【図14】



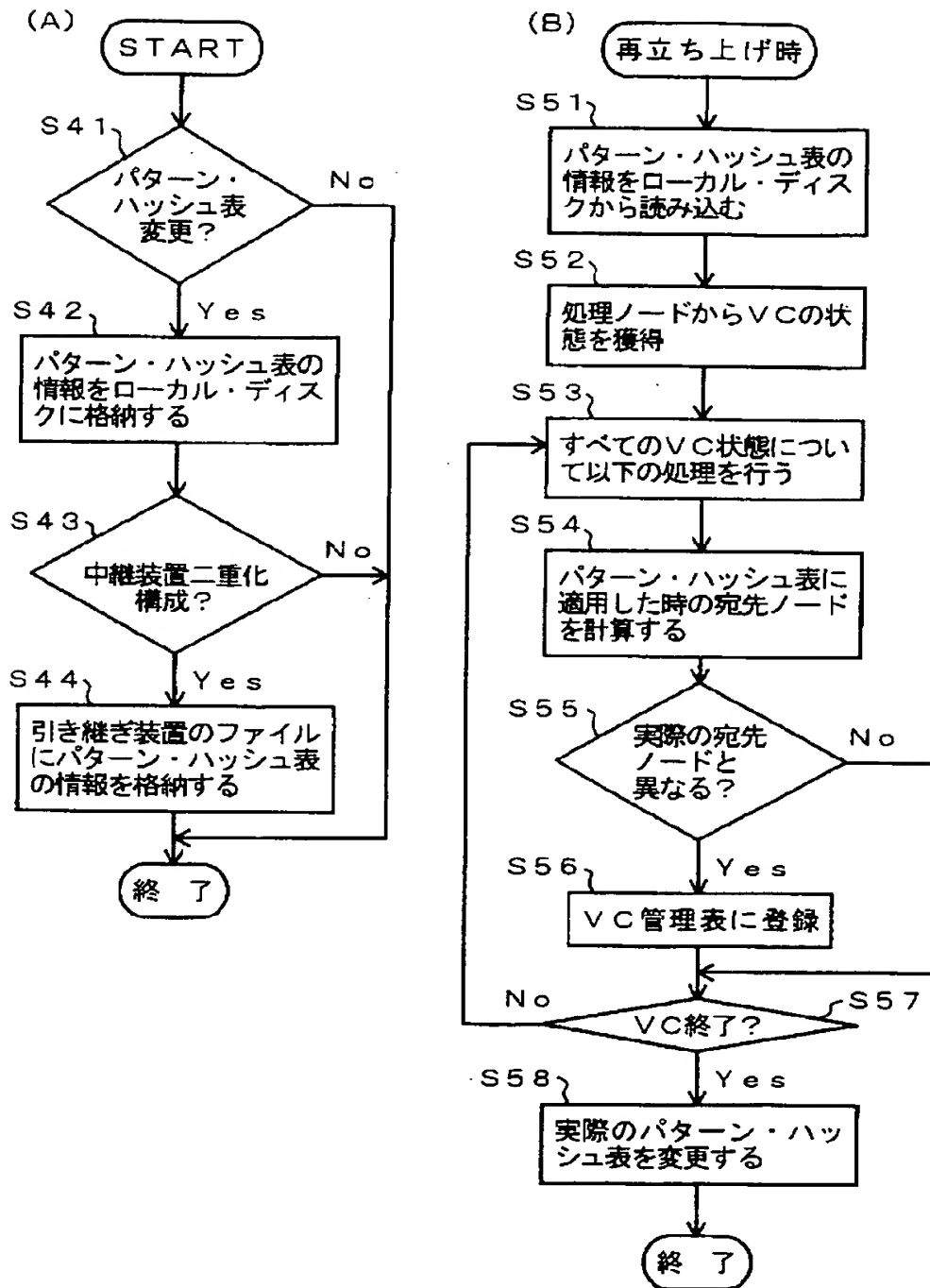
【図11】

VC管理表を用いる場合の packets 受信時の処理のフローチャート



【図12】

管理情報引き継ぎ処理のフローチャート



【図 16】

